

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 September 2003 (18.09.2003)

PCT

(10) International Publication Number  
WO 03/076575 A2

- (51) International Patent Classification<sup>7</sup>: C12N Sergei [UA/US]; 11 Stoney Point Ct., Germantown, MD 20876-5565 (US).
- (21) International Application Number: PCT/US03/06664
- (22) International Filing Date: 4 March 2003 (04.03.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/361,742 4 March 2002 (04.03.2002) US  
60/380,423 14 May 2002 (14.05.2002) US  
60/410,974 16 September 2002 (16.09.2002) US
- (71) Applicant (for all designated States except US): FIDELITY SYSTEMS, INC., et al. [US/US]; 7961 Cessna Avenue, Gaithersburg, MD 20879 (US).
- (71) Applicant (for US only): MALYKH, Andrei [RU/US]; 21504 Sun Garden Ct., Germantown, MD 20876-6942 (US).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): SLESAREV, Alexei, I. [RU/US]; 18406 Broad Leaf Road, Boyds, MD 20741-4320 (US). PAVLOV, Andrey [RU/US]; 12055 Suffolk Ter, Gaithersburg, MD 20878-2042 (US). PAVLOVA, Nadezhda [RU/US]; 12055 Suffolk Ter, Gaithersburg, MD 20878-2042 (US). KOZYAVKIN, (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: THE COMPLETE GENOME AND PROTEIN SEQUENCE OF THE HYPERTHERMOPHILE METHANOPYRUS KANDLERI AV19 AND MONOPHYLY OF ARCHAEAL METHANOGENS AND METHODS OF USE THEREOF

(57) Abstract: We have determined the complete 1,694,969 nucleotide sequence of the GC-rich genome of *Methanopyrus kandleri* using a novel approach. It is based on unlinking genomic DNA with the ThermoFidase version of *M. kandleri* topoisomerase V and cycle sequencing directed by 2'-modified oligonucleotides (Fimers). 3.3x sequencing redundancy was sufficient to assemble the genome with < 1 error per 40 kb. Using a combination of sequence database searches and coding potential prediction, 1692 protein-coding genes and 39 genes for structural RNAs were identified. *M. kandleri* proteins show an unusually high content of negatively charged amino acids, which might be an adaptation to its high intracellular salinity. Previous phylogenetic analysis of 16S RNA suggested that *M. kandleri* belonged to a very deep branch, close to the root of the archaeal tree. However, genome comparisons, using both trees constructed from concatenated alignments of ribosomal proteins and trees based on gene content, indicate that *M. kandleri* consistently groups with other archaeal methanogens. *M. kandleri* shares the set of genes implicated in methanogenesis and, in part, its operon organization with *Methanococcus jannaschii* and *Methanothermobacter thermoautotrophicus*. These findings indicate that archaeal methanogens are monophyletic. A distinctive feature of *M. kandleri* is the paucity of proteins involved in signaling and regulation of gene expression. Also, *M. kandleri* appears to have fewer genes acquired via lateral transfer than other archaea. These features might reflect the extreme habitat of this organism.

WO 03/076575 A2